

Faculdade de Saúde Pública

VOLUME 34
NÚMERO 5
OUTUBRO 2000
p. 547-59

Revista de Saúde Pública

Journal of Public Health

Avaliação de programas, serviços e tecnologias em saúde

Evaluation of health programs, services and technologies

Hillegonda Maria D Novaes

Departamento de Medicina Preventiva da Faculdade de Medicina da Universidade de São Paulo. São Paulo, SP, Brasil

Revisão

Review

Avaliação de programas, serviços e tecnologias em saúde*

Evaluation of health programs, services and technologies

Hillegonda Maria D Novaes

Departamento de Medicina Preventiva da Faculdade de Medicina da Universidade de São Paulo. São Paulo, SP, Brasil

Descritores

Avaliação[#]. Garantia da qualidade dos cuidados de saúde[#]. Avaliação de programas, métodos[#]. Serviços de saúde, avaliação[#]. Pesquisa sobre serviços de saúde[#]. Avaliação de tecnologias em saúde[#]. Planos e programas de saúde.

Keywords

Evaluation[#]. Quality assurance, health care[#]. Program evaluation, methods[#]. Health services, evaluation[#]. Health services research[#]. Evaluation of health technologies[#]. Health plans and programmes.

Resumo

A área de avaliação de programas, serviços e tecnologias em geral e na saúde, em particular, passa por um processo de expansão e diversificação conceitual e metodológica, bem como por uma crescente demanda para se constituir em instrumento de apoio às decisões necessárias à dinâmica dos sistemas e serviços de saúde, na implementação das políticas de saúde. Apoiando-se em uma revisão da literatura internacional especializada, e tomando por referência a década de 90, foram identificados os critérios nucleares que organizam os processos de avaliação, articulando-os com os recortes adotados pelas principais tipologias de avaliação, atualmente institucionalizadas nos países desenvolvidos, avaliação de programas, avaliação e garantia de qualidade em serviços e avaliação de tecnologias. A participação brasileira no desenvolvimento metodológico da área também é analisada.

Abstract

The field of program, services and technology evaluation in general, and in health care in particular, is going through an important growth and conceptual and methodological diversification. It is also the object of an increasing demand for its participation as an effective supportive instrument in the decision-making process, and a constant need in the dynamics of health systems and services. In this paper, based on literature review, nuclear criteria involved in the organization of all evaluation processes are identified, and articulated with the existing institutionalized evaluation practices in developed countries, that is, program evaluation, quality assessment and management and technology assessment. In conclusion, the incipient development of a methodological evaluation output in Brazil is analyzed.

INTRODUÇÃO

Condições de produção e de utilização de conhecimento: pensar e agir, saber e fazer, conhecimento e técnica, ciência e tecnologia.

A área de avaliação de programas, serviços e tecnologias, como campo produtor de saberes e práti-

cas, vivência de forma evidente e aguda as dificuldades a serem enfrentadas por aqueles que buscam produzi-la, na identificação das condições necessárias para a construção do conhecimento, produto de natureza abstrata, e daqueles que organizam a sua transformação em práticas, sempre concretas.

Não se trata da questão mais geral, ou seja, da pos-

Correspondência para/Correspondence to:
Hillegonda Maria Dutilh Novaes
Av. Dr. Arnaldo, 455
01246-903 São Paulo, SP, Brasil
E-mail: hidutilh@usp.br

*Parte da tese de livre docência apresentada à Faculdade de Medicina da Universidade de São Paulo, 1999.
Edição subvencionada pela Fapesp (nº 00/01601-8).
Recebido em 24/8/1999. Reapresentado em 21/6/2000. Aprovado em 23/8/2000.

sibilidade de produção de um conhecimento do mundo pelo homem, que pretende conter uma verdade, reconhecida como tal também fora do seu contexto imediato de produção, e que se traduz em temas como os da indução, dedução, comprovação, demonstração etc (Quine & Ullian,⁵⁰ 1970; Glymour,²⁵ 1992). Além dessa questão, colocada para toda atividade que se pretende científica e para a qual as diversas áreas têm buscado as soluções que lhes parecem mais adequadas, em determinados momentos dados (Altman,³ 1994; Chalmers,¹¹ 1994), trata-se do seguinte: quais as condições de utilização do conhecimento produzido? Em outras palavras, pode ele, e de que forma, ser inserido em processos históricos concretos que, além de terem delimitado as suas condições de produção, estarão também interferindo nas suas condições de utilização?

A abordagem desse problema passa necessariamente pela adesão a uma teoria ou explicação geral sobre o que faz o mundo se transformar, as coisas acontecerem, ao longo da história, apoiando-se em argumentos de natureza mágica, religiosa, positivista, materialista-histórica e outros. Em todas as teorias, o conhecimento produzido pelo homem sobre a realidade que se lhe apresenta, tem um sentido, uma razão de ser. No mundo moderno e contemporâneo ocidentalizado, em que passam a dominar as teorias racionalistas, o conhecimento, que na sua forma considerada superior ou mais elaborada é chamada de ciência, é considerado um instrumento essencial para a resolução dos problemas e para a construção de um mundo melhor. São fortes, ainda que diversas, as explicações de cunho teleológico para o desenvolvimento científico, bem como do uso da razão como uma razão instrumental, ou seja, a razão para alguma coisa e não a razão em si (Parrochia,⁴⁵ 1997). Dessa forma, apresenta-se como necessária uma proximidade, uma relação de consequência entre a ciência e a técnica, o pensar e o agir. Observa-se nesse processo uma “cientifização” da técnica e uma “tecnificação” da ciência, que assim, progressivamente, se interpenetram e se mostram capazes de solucionar todos os problemas, de qualquer natureza.

No entanto, a “cientifização” da arte (técnica) não é mais que uma concretização avançada de um velho sonho: tornar racional o mundo por meio daquele que, apreendendo-o, compreende-o e nele age como máxima causa eficiente. Fundando-se na racionalidade (científica), a explicação dos fenômenos e as soluções dela derivadas deparam-se com exigências centradas em valores socialmente tidos como fundamentais. Em outras palavras, ciência e tecnologia, além de terem finalidades utilitárias pre-

viamente definidas (o bem), devem responder positivamente a demandas mais amplas, como se delas decorressem os conflitos responsáveis pelas desigualdades e conflitos sociais (Novaes,⁴³ 1996). A repetida constatação da não-realização das proposições de máxima idealização, colocadas para a ciência e a técnica, tem sido usada como argumentação para duas atitudes opostas: (1) a afirmação de uma insuficiência “interna” da ciência e da técnica, ou seja, de natureza quantitativa, ou de intensidade, possível de ser superada, ou (2) da sua insuficiência “externa”, de natureza qualitativa, que as faz prometerem e serem responsabilizadas por algo que jamais poderá ser atingido (Portocarrero,⁴⁸ 1994). Esses posicionamentos, e suas variações, além de presentes nos debates filosóficos, ocupam atualmente um espaço importante nos ideários político-ideológicos que referenciam a prática política no mundo ocidental, com o predomínio do primeiro posicionamento, orientando as propostas para as políticas educacionais, científicas e tecnológicas (Witkowski,⁶⁰ 1991).

Essa questão pode ser recuperada e melhor compreendida no que diz respeito às suas repercussões para a prática científica e técnica cotidiana, quando sua presença é identificada na dinâmica de áreas de conhecimentos e práticas específicas. Para a análise do desenvolvimento da área de avaliação de programas, serviços e tecnologias em saúde, objeto do presente estudo, e tentando identificar sob que condições e em que contextos foi construída a articulação entre as condições de produção e de utilização do seu conhecimento, foi desenvolvida uma revisão da literatura produzida na década de 90, sobre esse tema, predominantemente em língua inglesa, com o objetivo de construir tipologias e categorias de análise mais gerais.

Não serão identificadas quais as modalidades mais adequadas ou melhores de avaliação, visando a construir uma compreensão sobre quais as condicionantes que incidiram sobre os diferentes desenvolvimentos na área. Essa posição é importante para a melhor compreensão da construção das avaliações em saúde, uma vez que ela toma como ponto de partida a produção de um conhecimento que contenha possibilidade de utilização ao assumir um conteúdo valorativo, o que necessariamente coloca esse tema no mundo da vida em sociedade.

Essa idéia de valor vai além do julgamento sobre o mérito do conhecimento como tal, mas pressupõe um posicionamento do avaliador e do avaliado sobre o conteúdo das conclusões, que incorpora e interage com o contexto no qual a avaliação se desenvolve.

AValiação em Saúde: Conceitos e Métodos para a Construção de Tipologias

As avaliações em saúde, ao se constituírem em uma área ainda em construção conceitual e metodológica, podem ser encontradas na literatura de forma muito diversificada (House,³¹ 1980). No presente estudo, desenvolve-se uma síntese das principais modalidades propostas, organizando-as, em um primeiro momento, em grandes categorias, segundo critérios considerados mais significativos para as questões aqui priorizadas, para em seguida inseri-las na nomenclatura mais frequentemente adotada. Os critérios apresentados foram selecionados a partir daqueles mencionados na literatura (Aguilar & Ander-Egg,² 1994; Hartz,³⁰ 1997; Stenzel,⁵⁸ 1996; Patton,⁴⁶ 1997) e discutidos em reuniões científicas. Esses critérios, que procuram dar conta das principais variáveis que orientam as decisões conceituais e metodológicas na construção dos processos de avaliação, são os seguintes: (1) objetivo da avaliação, (2) posição do avaliador, (3) enfoque priorizado, (4) metodologia predominante, (5) forma de utilização da informação produzida, (6) contexto da avaliação, (7) temporalidade da avaliação e (8) tipo de juízo formulado.

Para cada uma dessas variáveis, podem ser formuladas características alternativas que devem, no entanto, ser compreendidas como representando formas dominantes, mas não exclusivas, pela própria complexidade do objeto. Não existem, para as variáveis, modalidades melhores ou piores, ou seja, elas não têm um conteúdo valorativo em si, mas procuram apontar as alternativas conceituais e metodológicas existentes nos processos de avaliação, referenciadas por teorias e metodologias gerais aceitas por áreas de conhe-

cimento e práticas legitimadas. Um conteúdo valorativo será colocado quando for inadequada a utilização desses conceitos e metodologias isoladamente, ou ao se propor combinações contraditórias na análise de processos de avaliação concretos.

Os critérios ou variáveis selecionadas são:

- *Objetivo da avaliação*: priorização das condições de produção do conhecimento ou das condições de utilização do conhecimento (tomadas de decisão, aprimoramentos na gestão).
- *Posição do avaliador*: externo ou interno, por referência ao objeto avaliado.
- *Enfoque priorizado*: interno, de caracterização/compreensão de um contexto, ou externo, de quantificação/comparação de impactos de intervenções.
- *Metodologia predominante*: quantitativa ou qualitativa, situacional ou experimental/quasi-experimental.
- *Contexto da avaliação*: controlado ou natural.
- *Forma de utilização da informação produzida*: demonstração/comprovação ou informação, instrumentalização.
- *Tipo de juízo formulado*: comprovação/negação de hipóteses, recomendações ou normas.
- *Temporalidade da avaliação*: pontual, corrente, contínua.

Ainda que seja possível observar alguma variação nas combinações entre as características observadas para os critérios, de um modo geral, a adoção de uma posição para um critério coloca-se como condicionante para o enquadramento nos demais. Nesse sentido, é possível identificar três grandes tipos de avaliação, a partir da combinação entre as alternativas colocadas para cada critério: pesquisa de avaliação, investigação avaliativa, avaliação para decisão e avaliação para gestão (Tabela).

Tabela – Caracterização de tipos de avaliação segundo critérios selecionados.

Crítérios	Investigação avaliativa	Avaliação para decisão	Avaliação para gestão
Objetivo	Conhecimento	Tomada de decisão	Aprimoramentos
Posição do avaliador	Externo (interno)	Interno/externo	Interno/externo
Enfoque priorizado	Impactos	Caracterização/compreensão	Caracterização/quantificação
Metodologia dominante	Quantitativo (qualitativo) experimental/ quasi-experimental	Qualitativo e quantitativo situacional	Quantitativo e qualitativo situacional
Contexto	Controlado	Natural	Natural
Utilização da informação	Demonstração	Informação	Instrumentos para gestão
Juízo formulado em relação à Temporalidade	Hipóteses Pontual/replicado	Recomendações Corrente/pontual	Normas Integrado/contínuo

Na pesquisa de avaliação, o objetivo principal ou prioritário para o seu desenvolvimento é a produção de um conhecimento que seja reconhecido como tal pela comunidade científica, ao qual está vinculado, conhecimento que servirá como fator orientador de decisão quando se colocarem questões como viabilidade, disponibilidade de tempo e de recursos e demandas externas à pesquisa (Greene,²⁷ 1994; Mohr,³⁸ 1995; Minayo,³⁷ 1992; Rossi & Freeman,⁵³ 1993). Essas pesquisas geralmente são desenvolvidas sob a coordenação de instituições acadêmicas, por iniciativa delas próprias, ou a partir de solicitações de instâncias públicas com elevado poder sobre a implementação ou reorientação de macro políticas (Ministérios ou Secretarias Estaduais, Legislativo e outros órgãos), nos quais, à distância dos avaliadores, ou de pelo menos parte da equipe responsável pelas principais decisões (não implicando a não-participação de avaliadores internos no processo), dada a valorização da máxima objetividade possível, é considerada uma condição necessária ao adequado desenvolvimento da investigação.

As principais questões que orientam as pesquisas de avaliação, ou seja, o enfoque priorizado para a formulação dos objetivos a serem atingidos, dizem respeito preponderantemente à identificação de impactos obtidos pelas ações a serem avaliadas. Ainda que passíveis de serem formulados de forma mais direta, ou mais mediatizada, o estabelecimento de nexos entre uma ação, ou ações, e determinadas alterações observadas na realidade (indivíduos, populações, práticas, ambiente etc), orienta fortemente as questões formuladas. Ou seja, pode-se considerar como característica importante nas pesquisas de avaliação a identificação de relações de causalidade, ainda que as concepções que as orientam possam ser diversificadas (positivistas, “verificacionistas”, complexas e outras) (Rothman,⁵² 1988; Chen,¹² 1990).

A metodologia – ou seja, as formas propostas para a aproximação e a apreensão do objeto – dominante nas pesquisas de avaliação ainda é a de natureza quantitativa, mas com crescente participação das formas de natureza qualitativa, ambas porém orientadas para a busca da objetividade e de uma possibilidade de generalização do observado. Isso implica o estabelecimento de um certo controle sobre o contexto em que se desenvolve a pesquisa e na proposição de desenhos gerais metodológicos experimentais ou, mais freqüentemente, quasi-experimentais (buscando alcançar um equilíbrio satisfatório entre o controle da subjetividade e a garantia de uma objetividade, a artificialidade, a viabilidade e a ética, dilemas sempre presentes nas investigações de um modo geral, e principalmente nas de avaliação) (Miles e Huberman,³⁶ 1994; Silva & Formigli,⁵⁷ 1994; Patton,⁴⁶ 1997).

Diante dessas condicionantes, a informação produzida, que se deseja ser um conhecimento, deverá ser utilizada para a demonstração dos pressupostos iniciais que orientaram todo o processo, e o juízo formulado como resultado da pesquisas de avaliação estará relacionado a esses mesmos pressupostos, ou hipóteses. Quanto à temporalidade dessas investigações, considerando-se que elas, em geral, têm por objetivo responder a questões cuja necessidade de resposta, ainda que genericamente previsível, não foi planejada com antecedência, no que se refere ao momento e às formulações específicas, e dado o seu elevado custo, são elas habitualmente de natureza pontual, podendo eventualmente ser replicadas, temporal ou espacialmente, para efeito comparativo.

Na avaliação para a decisão, o objetivo dominante é ela se constituir em um elemento efetivamente capaz de participar de processos de tomada de decisão, ou seja, que produza respostas para perguntas colocadas por aqueles que vivenciam o objeto avaliado, sendo esta capacidade mais importante, se a escolha se fizer necessária, do que o cientificismo do conhecimento produzido. Sendo assim, o avaliador interno passa a ter uma posição decisiva no desenrolar do processo, ainda que avaliadores externos geralmente também façam parte da equipe coordenadora da avaliação. E o enfoque priorizado, aquele que orienta os principais objetivos, é o do reconhecimento do objeto, na profundidade necessária para a sua adequada compreensão, identificação dos problemas e alternativas de equacionamento possíveis (Guba & Lincoln,²⁸ 1989; Greene,²⁷ 1994).

Tendo em vista a natureza complexa dos objetos colocados para essas avaliações, são utilizadas metodologias de natureza qualitativa e quantitativa, de natureza situacional, com desenhos relativamente abertos e específicos para cada tipo de aproximação e apreensão da realidade buscada (como, por exemplo, os estudos de caso), em um contexto que se deseja manter natural e cujos resultados deverão ser articulados no momento da análise geral, ao se tomar por referência uma teoria geral que permita a sua integração (Chen,¹² 1990).

Esses resultados constituem informações geralmente reconhecidas como melhor construídas, mais sistematizadas, do que aquelas inicialmente disponíveis sobre o objeto avaliado, mas cujo valor final não é intrínseco e será medido por sua capacidade de se transformar em recomendações positivas, isso é, capazes de contribuir para a solução dos problemas identificados (pelos demandantes da avaliação, ou identificados ao longo do processo) (Patton,⁴⁶ 1997). Esse tipo de avaliação, ainda que realizado pontualmente,

isso é, como processo especialmente desenvolvido para esse fim, é de natureza corrente, dado que geralmente previsto para acontecer quando da proposição e implantação do objeto avaliado.

A avaliação para gestão tem como objetivo principal a produção da informação que contribua para o aprimoramento do objeto avaliado. Não se trata da busca de fundamentos ou justificativas nem redirecionamentos, mas o melhoramento de uma condição dada. Nessa medida, a presença do avaliador interno é condição necessária e é nos seus termos que se dá a participação do avaliador externo, geralmente presente.

O enfoque priorizado, explicitado nos objetivos, é o da caracterização de uma condição e a sua tradução em medidas que possam ser quantificadas e replicadas. São dominantes as metodologias de natureza quantitativa (podendo estar também presentes instrumentos com abordagem qualitativa, desde que passíveis de uso padronizado), observacionais, desenvolvidas em contexto natural ou até mesmo de rotina administrativa, em que os desenhos gerais das avaliações são construídos pela agregação de medidas representativas de uma totalidade. A informação produzida é prioritariamente voltada para o desenvolvimento e aprimoramento de indicadores, como partes que representam um todo, e o resultado desejado para essa avaliação é a proposição de critérios ou normas a serem incorporadas na utilização desses indicadores no desempenho rotineiro do objeto avaliado, contribuindo para a garantia do seu bom funcionamento (Donabedian,¹⁸ 1986; Scrivens,⁵⁵ 1995; Schoenbaum et al,⁵⁴ 1995).

AValiação em Saúde: OBJETOS PARA O CONHECIMENTO E PRÁTICAS SOCIAIS

As avaliações de programas, serviços e tecnologias de saúde inserem-se, com graus variáveis de aderência, em uma das tipologias propostas, dependendo do contexto no qual se desenvolvem. Não se trata de identificar tipologias avaliativas *a priori*, melhores ou piores, como instrumentos ativos de construção social, mas apontar a articulação necessária entre objetos, tipologias e contextos, nos processos de avaliação.

No desenvolvimento das avaliações em saúde como práticas institucionalizadas, em conjunturas diversas, foram sendo construídas especializações e segmentações, com inserções e perfis profissionais próprios. Com atividades de avaliação em programas, serviços e tecnologias de saúde, podem atualmente ser identificadas diferentes “comunidades” de profissionais, inseridos em setores diversos e mais identificados com referenciais político-ideológicos e teórico-metodológicos específicos, geralmente vinculados a

associações científicas e/ou profissionais que os representam. Entre as mais atuantes serão destacadas aquelas que desenvolvem práticas denominadas de “avaliação tecnológica em saúde”, “avaliação de programas” e “avaliação, gestão e garantia de qualidade”. Nos países onde estão presentes, na dependência de fatores como as características da organização da atenção à saúde, formação e “cultura” profissional, entre outras, essas “comunidades” se inserem, de maneiras diversificadas, na saúde pública, na administração e na gestão ou na medicina.

Avaliação tecnológica em saúde

Na sua definição mais abrangente, a avaliação tecnológica em saúde é aquela que toma como sua unidade de análise, ou ponto de partida, uma tecnologia, de produto ou de processo, passível de ser caracterizada na sua dimensão temporal e espacial (que, onde, como, quando, para quem, para quê). Considerando-se a complexidade atual da atenção à saúde, o estudo de tecnologias específicas pressupõe sempre uma circunscrição do objeto relativamente arbitrária, que é justificado pela possibilidade de um aprofundamento maior no reconhecimento daquela tecnologia. Não significando isso, necessariamente, uma limitação da perspectiva da análise, que pode estar sendo desenvolvida a partir do ponto de vista dos indivíduos nos quais a tecnologia é utilizada, com ênfase nas dimensões biológicas, até aquele dos macrocontextos políticos e econômicos em que ela se insere (Goodman,²⁶ 1992; Banta & Luce,⁵ 1993). Contudo, nos recortes analíticos mais conjunturais, em que as tecnologias são tomadas como parte de processos econômicos e sociais historicamente determinados, há uma valorização do conhecimento do específico, do técnico, como caminho necessário para a constituição de um geral.

A avaliação tecnológica em saúde, como atividade institucionalizada, desenvolveu-se nos países centrais a partir dos anos 70, associada à expansão da atenção à saúde e ao desenvolvimento científico e tecnológico em saúde que levou à introdução de muitas novas tecnologias, com implicações políticas, econômicas, administrativas e sanitárias importantes. É praticada atualmente, de forma predominante, em comissões ou serviços de assessoria em instâncias diversas do setor público (executivo, legislativo ou universitário), mas também por associações profissionais e empresas privadas (Battista & Hodge,⁸ 1995; Perry, Gardner & Thamer,⁴⁷ 1997; Luce & Brown,³³ 1995; Banta & Perry,⁷ 1997).

Uma das formas propostas de avaliação tecnológica busca responder à necessidade de conhecimentos

aprofundados sobre algumas tecnologias, no que diz respeito às suas dimensões técnicas (a segurança e efetividade de uma tecnologia, ou seja, a capacidade de produzir o resultado para o qual ela foi desenvolvida, sem causar danos, em condições normais de utilização nos serviços de saúde) e econômicas (estudos de custo-efetividade, custo-utilidade ou custo-benefício), quando então se associam às avaliações econômicas (Drummond et al,²¹ 1997; Elixhauser et al,²² 1998).

As avaliações de tecnologias dessa natureza, dadas as suas condicionantes, constituem pesquisas de avaliação que se encaixam na tipologia anteriormente apresentada. Além disso, têm sido desenvolvidas para o estudo de novas tecnologias, incorporadas há pouco tempo nos serviços de saúde, ou daquelas que passaram a ser utilizadas em contextos diversos de quando inicialmente incorporadas (por exemplo, do ambiente hospitalar para o ambiente ambulatorial, para quadros clínicos mais diversificados, em associação a outras tecnologias entre outras), ou ainda, daquelas de uso corrente, mas para as quais existem dúvidas sobre sua efetividade atual.

Mais recentemente, os profissionais envolvidos nas avaliações tecnológicas em saúde voltadas para o estudo das efetividades têm se aproximado das áreas da epidemiologia e clínica que estudam as eficácias diagnósticas e terapêuticas dos procedimentos utilizados na prática médica e das atividades relacionadas à “Medicina Baseada em Evidências” (AHCPR,¹ 1997), por meio de um reconhecimento mútuo das relações de contigüidade (ainda que não de continuidade) existentes entre as respectivas práticas (Novaes,⁴² 1996). Além dessas, compõem também esse tipo de avaliação tecnológica as pesquisas em engenharia biomédica e engenharia clínica, que se dedicam à avaliação das condições de produção, operação e manutenção dos equipamentos e materiais de uso especializado na área da saúde (Goodman,²⁶ 1992).

As atividades relacionadas ao desenvolvimento de consensos e “guidelines” para a utilização de tecnologias na área de saúde, mais intensamente presentes nos anos 90, e que têm sido propostas e coordenadas tanto por órgãos públicos executivos centrais quanto por associações profissionais ou mesmo serviços, caracterizam-se como avaliações tecnológicas do tipo “avaliação para decisão”, ao terem por objetivo a utilização do conhecimento científico atual na boa prática profissional, propondo recomendações capazes de servir de apoio nos processos de decisão (administrativos, gerenciais, clínicos) (Banta et al,⁶ 1995; Panerai & Mohr,⁴⁴ 1989).

Diferenciam-se dos critérios e das normas dos anos 70, utilizados nos serviços de atenção à saúde, pouco

preocupados com a sua fundamentação em sínteses do conhecimento existente sobre uma determinada questão, conhecimento freqüentemente não concordante nas suas conclusões e para cuja sistematização necessariamente terão de ser utilizados critérios de validação que precisam ser aceitos como legítimos pela comunidade que irá utilizar seus resultados, bem como a sua tradução em recomendações à prática. Os critérios e normas dos anos 90, presentes e utilizados nos processos de garantia de qualidade, a serem discutidos mais adiante, buscam fortalecer sua legitimidade, pelo menos como discurso, na articulação com os consensos e “guidelines” pertinentes e na sua inserção em processos de trabalho concretos.

Finalmente, têm sido incluídos no universo da avaliação tecnológica em saúde, apresentando características mais acentuadas de pesquisa e de produção de conhecimento, os estudos relativos às políticas científicas e tecnológicas para a área de equipamentos e medicamentos, além daqueles que tomam como sua questão principal as dimensões tecnológicas, mas os desenvolvem na perspectiva dos estudos de “práticas e políticas de saúde”, mais próximos das pesquisas propriamente sociológicas ou históricas.

Avaliação de programas em saúde

As avaliações de programas em saúde têm como foco de análise os programas (tomados aqui no sentido em que é utilizado na literatura americana), como processos complexos de organização de práticas voltadas para objetivos especificados. Portanto, são considerados programas, tanto aquelas propostas voltadas para a realização de um macro objetivo, como a implantação de formas de atenção para populações específicas (atenção domiciliar para idosos, vacinação e procedimentos de “screening” para grupos de risco para problemas de saúde determinados, propostas terapêuticas para doenças e doentes priorizados etc.) e que envolvem instituições, serviços e profissionais diversos, como as atividades desenvolvidas em serviços de saúde, que têm por objetivo prestar um determinado tipo de atendimento para uma dada clientela (Shortell & Richardson,⁵⁶ 1978). No Brasil, de um modo geral, apenas os procedimentos da primeira alternativa são considerados programas, mas já se observa também aqui uma maior variabilidade no seu uso, o que faz com que toda avaliação de programa deva se iniciar com uma clara explicitação dos contornos do objeto a ser avaliado: propostas, ações programáticas, atividades gerenciais, serviços, procedimentos, profissionais etc.

No que diz respeito ao objeto colocado para a avaliação e às alternativas de apreensão do mesmo, a

avaliação de programas mostra, de um lado, interfaces com a avaliação tecnológica e a avaliação de qualidade (Donabedian,¹⁶ 1988; Goodman,²⁶ 1992), e, de outro, com a ciência social e política e o planejamento (Rivera,⁵¹ 1996), pois objetivos programáticos realizam-se por sistemas, serviços e procedimentos concretos. Ao mesmo tempo, fazem parte da conjuntura política, econômica e social mais ampla, conjunturas essas que não são passíveis de abordagens racionalizadoras relativamente circunscritas, como são as avaliações.

O dilema da articulação entre a produção de conhecimento e a sua utilização, ou entre a prática científica e técnica e a prática política, e as alternativas desenvolvidas para o seu equacionamento, fazem parte da história da Avaliação de Programas como prática institucional (Cabatoff,¹⁰ 1996). Essa se desenvolveu principalmente nos Estados Unidos, a partir dos anos 60, como parte dos grandes programas de intervenção social propostos na época, resultando no surgimento de uma categoria profissional, os avaliadores de programas, que estão vinculados a instituições públicas ou privadas, ou organizados em empresas prestadoras desse tipo de serviço, e seus reflexos podem ser observados nos tipos de avaliação de programas (em geral e da saúde) atualmente propostos na literatura.

Na fase inicial, a sua dimensão de instrumento para intervenção articulava-se com a atuação de um Estado que devia garantir um bem-estar social, porém de forma eficiente. Nos anos 90, essa dimensão, refletindo as grandes alterações ocorridas na atuação do poder público e nas teorias de ação social, aparece quando é colocada a possibilidade do “empowerment” dos agentes do programa (profissionais e população) pela avaliação (Patton,⁴⁶ 1997), ou seja, de uma perspectiva de dever institucional passou-se para a mobilização individual.

Dado o seu caráter sempre conjuntural e de contexto, uma das formas frequentes de criar categorias para as avaliações de programas é pela especificação da fase (ou fases) do programa que está sendo avaliado (de desenvolvimento e proposição, também denominado de estratégico, de implantação e de resultados e impacto) e pelo detalhamento da questão, ou questões, que orientam a avaliação ou o(s) foco(s) a serem privilegiados (descritivo, causal, comparativo, prospectivo, econômico, mobilização social, monitoramento entre outros), gerando longas listas de tipos de avaliação, que mais confundem do que auxiliam na sua compreensão (Contandriopoulos et al,¹³ 1997).

Esse nominalismo, que atende principalmente a uma necessidade de oferecer sempre novos produ-

tos para o mercado, por vezes dificulta a percepção de que se organizam as avaliações de programas, de um certo número de teorias de ação social, de formas de produzir conhecimento e de alternativas metodológicas, como alguns dos principais debates no interior do campo (“fourth generation evaluation”, “theory-driven evaluation”, “utilization-focused evaluation”) (Guba & Lincoln,²⁸ 1989; Chen,¹² 1990; Patton,⁴⁶ 1997) se propõem a articulá-los e transformá-los em propostas práticas, orientados por opções que são também políticas.

Retomando as tipologias anteriormente apresentadas, é possível identificar nas avaliações de programas a presença da pesquisa de avaliação, da avaliação para decisão e da avaliação para gestão, de forma isolada ou concomitante. A pesquisa de avaliação desenvolve-se mais frequentemente, quando o seu objetivo é conhecer o impacto de um programa, sendo esse medido pelos resultados observados sobre condições específicas da população (mortalidade, morbidade, indicadores de qualidade de vida entre outras), selecionados a partir de relações de causalidade identificadas entre condições de saúde e formas de intervenção. Nesse caso, predominam as metodologias quantitativas, articuladas a alguma forma de controle sobre os fatores “de confusão”, e o conhecimento produzido deve ter validade interna e externa, ter confiabilidade e ser objetivo, para poder ser reconhecido como científico, em comunidades específicas (Mohr,³⁸ 1995; Yates,⁶² 1996).

No entanto, outros tipos de impacto têm sido medidos, e o intenso desenvolvimento das metodologias qualitativas, tanto da sua fundamentação teórica quanto das aplicações práticas, nos últimos anos (Denzin & Lincoln,¹⁴ 1994), está presente também nas avaliações de programas para esses aspectos. Quando o objetivo da análise desses impactos é a construção de um conhecimento passível de uma generalização, trata-se também de uma pesquisa de avaliação, e observa-se nesses estudos uma preocupação com a sua regularidade, objetividade, confiabilidade (“responsiveness”, “objectivity” e “trustworthiness”), com inserção em um referencial teórico mais geral (Chen,¹² 1990).

Também as avaliações estratégicas e de implantação podem constituir-se em pesquisa de avaliação, mas o fazem menos frequentemente, principalmente as segundas, pela própria dificuldade que seus recortes oferecem ao enquadramento teórico-metodológico mais rígido que a prática científica exige, tal como é concebida, por não estarem no geral orientadas para a demonstração de hipóteses explicativas previamente desenvolvidas.

As avaliações de implantação e estratégicas mostram-se mais aderidas às avaliações para decisão, pois as questões formuladas necessitam de uma abordagem mais holística, priorizando, quase sempre, o conhecimento e a compreensão da complexidade e das dinâmicas gerais. Busca responder a questões colocadas para aquele programa específico, e cujas respostas serão consideradas adequadas pela sua capacidade de resolver os problemas identificados, quanto ao mérito e melhorias possíveis, pelos grupos de interesse envolvidos ("stakeholders"), freqüentemente numerosos, e com perspectivas em geral não concordantes.

Nessas abordagens, é possível observar menor interesse por questões etiológicas ou de causalidade de referenciais explicativas mais gerais (o que tem sido chamado de "black box approach") e a priorização de sensibilidade e empatia maiores para com a "vida como ela é", a ser abordada por metodologias "construtivistas", apoiadas em teorias "relativistas" do conhecimento (essa é uma das formas em que se fez presente no campo das avaliações um debate muito mais amplo e bastante complexo) (Guba & Lincoln,²⁸ 1989; Bernstein,⁹ 1983).

No final dos anos 80, essa polarização entre o que se pode denominar de "ser cientista" e "ser solidário" mostrava-se de forma muito aguda na literatura, mas em publicações mais recentes é possível observar a busca de compromissos possíveis entre eles. De um lado, isso se traduz em uma abertura maior para a aceitação tanto de metodologias quantitativas quanto qualitativas em avaliações específicas, reconhecendo-as como possuidoras de especificidades próprias e complementares para a construção de um conhecimento e, por outro, há o reconhecimento da inevitável presença de algum referencial teórico ou conhecimento prévio, mesmo quando não explícito, e da interferência sobre o "natural" que qualquer metodologia provoca (Chen,¹² 1990; Patton,⁴⁶ 1997).

Observa-se também a presença de preocupação, quando da discussão das avaliações de tipo para decisão e para gestão, para com a diferenciação entre a aproximação "intuitiva" da realidade, a inquirição sistemática ("systematic inquiry", que é a que se aplicaria a elas) e a pesquisa, velhos e eternos dilemas da busca da verdade e da utilidade, que mesmo quando declarados ausentes estão presentes.

As avaliações para gestão podem ser encontradas nas avaliações de implantação e de impacto, basicamente, quando o objetivo é o monitoramento de algumas conclusões desenvolvidas por pesquisas de avaliação ou avaliações de decisão, ao se prever algum tipo de continuidade para a avaliação realizada.

Avaliação, gestão e garantia de qualidade em saúde

A idéia de qualidade está presente em todos os tipos de avaliação, uma vez que têm como característica nuclear o estabelecimento de um juízo, a atribuição de um valor a alguma coisa que, quando positivo, significa ter qualidade, na acepção atual do termo. Ou seja, uma avaliação tecnológica ter melhor custo-efetividade quer também dizer ter mais qualidade; desse ponto de vista, ou, na avaliação de um programa, o estabelecimento do mérito, por referência a determinados parâmetros, remete ao estabelecimento de uma qualidade. No entanto, na "avaliação, gestão e garantia de qualidade", esse conceito deixa de estar implícito, passando a ter uma função teórica e prática central, ao se constituir em um ponto de partida (e não mais de chegada) para todo o processo (Donabedian,¹⁷ 1988).

Nessa perspectiva, a avaliação de uma tecnologia ou programa em que, de início, há o estabelecimento de uma qualidade que se constituirá no parâmetro de referência, a partir do qual serão construídos os instrumentos a serem utilizados na avaliação, passa a ser uma avaliação de qualidade. Contudo, são os serviços que mais freqüentemente têm se constituído em objeto para a avaliação de qualidade. Esse recorte preferencial pode ser melhor compreendido quando as origens e vinculações institucionais da avaliação de qualidade, bem como a sua estreita articulação com a gestão e garantia de qualidade, são identificados.

É possível considerar que é a partir do desenvolvimento econômico capitalista e industrial que o valor (de troca) de um produto passa a depender também do oferecimento de uma "garantia" de estabilidade no seu valor de uso, ou seja, de uma determinada qualidade. Os mecanismos para garantirem essa qualidade foram se modificando, acompanhando as transformações nos processos de produção industrial: a excelência dos agentes produtores da mercadoria, os sistemas de controle e fiscalização dos produtos, os sistemas de garantia da qualidade (normas e padrões) e a gestão ou controle total da qualidade (Donabedian,²⁰ 1996).

A partir dos anos 60, os serviços, por sua própria expansão, passaram a ser pensados também como produtos e, portanto, passíveis de uma padronização na qualidade, atualmente um atributo importante na sua valorização no mercado. Além da sua importância no setor privado, a qualidade ganhou destaque também nos serviços prestados no setor público, como parte dos direitos sociais da população e pelos movimentos de consumidores.

Os serviços de saúde, um dos setores de maior expansão no período, foram igualmente envolvidos nesses movimentos, mais intensamente no setor privado, mas também no setor público, quando então a avaliação da qualidade, frequentemente associada à gestão e à garantia de qualidade, passou a ser desenvolvida com alguma regularidade por profissionais especializados, em particular nos Estados Unidos.

Ainda que interligados, tanto a avaliação quanto a garantia e a gestão da qualidade têm seus referenciais teóricos e metodológicos específicos. Para a avaliação de qualidade, o autor mais importante é Donabedian, cujos primeiros trabalhos sobre o tema datam dos anos 60, mas que ganhou maior notoriedade mundial a partir do final dos anos 80.^{15,20} Seu ponto de partida é a definição do que estará sendo considerado qualidade, pois essa não se constitui em um atributo abstrato, mas deve ser construída em cada avaliação por meio dos, por ele denominados, sete pilares da qualidade: eficácia, efetividade, eficiência, otimização, aceitabilidade, legitimidade e equidade (Donabedian,¹⁸ 1990).

Cada uma dessas dimensões deverá ser especificada, para dados contextos, pela identificação dos critérios mais adequados para aquela situação, dando conta das suas particularidades e, ao mesmo tempo, mantendo uma coerência entre si, pois apresentam uma interdependência e um movimento que vai do mais específico ao mais geral. Isso é, as dimensões consideradas mais técnicas condicionam as dimensões “interpessoais”, mas ao mesmo tempo dependem delas para realizarem seu potencial. Além disso, os critérios e padrões podem estabelecer níveis mínimos ou máximos de qualidade e se orientar segundo as prioridades dos prestadores dos serviços, da visão dos usuários ou das prioridades definidas pela sociedade. São opções que deverão ser definidas para cada avaliação¹⁴ (Donabedian,¹⁷ 1988).

Na dependência dos objetivos a serem alcançados pela avaliação da qualidade do serviço, ela poderá constituir-se em pesquisa de avaliação, avaliação para decisão ou avaliação para gestão. Os dados que compõem os instrumentos utilizados para a apreensão do objeto poderão dizer respeito às condições estruturais, ou seja, aquelas condições que constroem e condicionam o universo de práticas a ser avaliado (recursos financeiros, humanos, físicos, organizacionais etc.), os dados relativos aos processos que definem essas práticas (consultas, internações, cirurgias, exames etc.) e as medidas que permitam a mensuração de resultados das mesmas (mortalidade, morbidade, medidas de qualidade de vida, notificações, tempo de internação etc.).

As avaliações de qualidade mais conhecidas de Donabedian dizem respeito à atenção médica, com ênfase na efetividade, eficiência e otimização dos procedimentos e serviços, seguindo o padrão das pesquisas de avaliação, que implicam geralmente a utilização de variáveis estruturais, de processo e de resultado, ao estabelecer relações de causalidade. No entanto, o seu referencial teórico também tem sido utilizado para outros profissionais, modalidades de atenção à saúde e tipos de avaliação, com a utilização de metodologias quantitativas e qualitativas diversas.

Contudo, é possível afirmar que há um maior aprofundamento nesse referencial teórico das condições de produção do conhecimento (seja ele científico ou não) do que das suas condições de utilização, tomadas como quase implícitas. Por exemplo, constitui-se em um processo tipicamente “donabediano” a transformação de “guidelines” construídas a partir de estudos clínicos, avaliações tecnológicas e avaliações de programas, em critérios e padrões de qualidade para um determinado serviço; esses serão validados por uma pesquisa de avaliação e depois avaliados como instrumentos de monitoramento em uma avaliação para gestão, mostrando-se adequados ou não para compor um programa de garantia de qualidade em um serviço.

Para cada um desses momentos, os objetivos e as metodologias que poderão viabilizá-los devem estar claramente definidos, não devendo ser valorizadas aquelas dimensões que dizem respeito às condições concretas dos serviços e que poderão interferir na incorporação desses conhecimentos. Observa-se nos textos de Donabedian¹⁹ (1992) uma certa desconsideração para com as condições operacionais, administrativas e de cultura organizacional.

Outro referencial teórico importante de problematização da questão da qualidade em saúde, e que valoriza esses aspectos, é a vertente mais fortemente articulada à gestão propriamente dita da qualidade, conhecida como Gestão da Qualidade Total – GQT, quando então a avaliação é menos problematizada. Dessa forma, a fundamentação teórica das práticas e a sua correta tradução em normas, critérios e padrões é considerada como geralmente já disponível; sendo decisivos para a existência da qualidade como parte do cotidiano dos serviços, o desenvolvimento de mecanismos que promovam a adesão de tudo, e todos aos objetivos, definidos como prioritários, estabelecidos da forma mais participativa possível.

Esse é um campo em grande expansão, com muitos autores e muitas publicações, sendo considerados como os principais teóricos do movimento Ishikawa³²

(1993), para o modelo japonês de gestão da qualidade, e Juran & Deming (citados por Nogueira,³⁹ 1994) na perspectiva dos países desenvolvidos ocidentalizados. Para a área da saúde, a referência internacional é Berwick (Nogueira,³⁹ 1994) e a perspectiva adotada é fundamentalmente a do setor privado, sendo o cliente o principal elemento definidor da qualidade a ser construída.

Na gestão da qualidade, o tipo de avaliação desenvolvido é a avaliação para decisão, podendo também ser complementado com uma avaliação para gestão, com as características gerais apontadas anteriormente e com metodologias de apreensão da realidade muito próximas daquelas adotadas pelo planejamento estratégico-situacional, apesar das significativas diferenças no seu referencial teórico (Rivera,⁵¹ 1996).

Uma forma específica de avaliação, gestão e garantia de qualidade que tem se mostrado importante na área da saúde é o desenvolvimento de “sistemas de acreditação”, basicamente para hospitais, mas com recentes propostas de sua extensão para redes de serviços de saúde (Scrivens,⁵⁵ 1995); articula esses três momentos, constituindo-se em uma forma combinada de avaliação para decisão e avaliação para gestão.

O país onde essa proposta se desenvolveu originalmente foi os Estados Unidos, ainda nos anos 20 e 30, por iniciativa de associações profissionais médicas, que procuravam desenvolver instrumentos que lhes garantissem condições mínimas de atuação profissional, face aos interesses dos proprietários dos hospitais privados. Tratava-se, portanto, de uma modalidade de defesa profissional, fortemente inserida na lógica geral de atuação profissional e de organização do sistema de atenção à saúde americana. Em períodos mais recentes, passou a ser considerado também um instrumento importante para a inserção institucional no mercado pelos hospitais privados, bem como um mecanismo de padronização e garantia de qualidade básica pelos gestores do setor público.

Dessa forma, a atual “Joint Commission on Accreditation of Healthcare Organizations” (JCAHO), criada pela American College of Surgeons, American College of Physicians e American Hospitals Association, além de algumas outras associações menores, é hoje responsável pela realização de um processo de avaliação dos hospitais e outros serviços de saúde, utilizando-se de uma metodologia padronizada, incluindo questionários, entrevistas e observações, que resulta em uma classificação da instituição avaliada, com aprovação total, parcial, condicionada ou rejeição. O conceito obtido nessa avaliação e a participação no sistema de monitoramento posterior

constituem um dos mais importantes instrumentos formais de garantia de qualidade para os hospitais americanos e parte ativa na negociação de acordos financeiros, busca de clientela e acordos de trabalho com os profissionais.

Mais recentemente, Austrália e Canadá têm desenvolvido propostas semelhantes, em sistemas de saúde com lógicas bastante diversas, em que há uma grande participação do Estado na organização da atenção à saúde, mas nos quais essa característica de delegação da avaliação das instituições hospitalares a uma organização externa, fora do setor público, tem se mostrado aceitável e até desejável. Isso já não acontece na Grã-Bretanha, um país que apesar de também ser anglo-saxão (pois em países com cultura latina esse tipo de prática é considerado ainda mais difícil de ser aceito) tem uma tradição de gestão pública, de um lado, e de autonomia profissional, de outro, diversa.

AVALIAÇÃO EM SAÚDE NO BRASIL

O eixo da análise desenvolvida no presente estudo procurou, ao identificar tipologias e modalidades de práticas na avaliação à saúde, localizá-las também em momentos e contextos históricos específicos, por considerá-los parte do seu processo de desenvolvimento. Nesse sentido, é importante reconhecer que a inserção das propostas e práticas de avaliação em saúde em contextos diversos daqueles nos quais foram gerados, ao serem “idéias fora de lugar”, necessitam (Novaes,⁴¹ 1979) de uma reinterpretação e resignificação para poderem fazer sentido.

A difusão das metodologias de avaliação em saúde para os países em desenvolvimento é realizada de diversas maneiras. Nos países com maior intercâmbio científico com os países centrais, ela se dá por meio da literatura científica, congressos e reuniões ou programas de pós-graduação. Além disso, diversas organizações internacionais, tanto de saúde (World Health Organization, Organização Panamericana de Saúde, United Nations Children's Fund) como financeiras (Banco Mundial), estimulam a sua utilização em projetos de pesquisa com propostas metodológicas estruturadas em “pacotes”, sendo a sua presença dominante nos países com pouca densidade científica local.^{59,61}

A justificativa apresentada para essas propostas fechadas e, geralmente, simplificadas, é a grande utilidade dos processos de avaliação para a adequada identificação das prioridades em saúde, a necessidade de um aprendizado rápido e de adaptação a contextos com recursos humanos menos qualificados e em menor número, recursos financeiros restritos, além de poucas informações previamente disponíveis. Alte-

ram-se, freqüentemente, características consideradas centrais para os processos de avaliação, ou seja, aqueles procedimentos que buscam fazer com que as avaliações articulem de forma efetiva o conhecer e o fazer, o que sempre pressupõe uma grande “personalização” dos processos.

Não foi tratada em profundidade, no presente estudo, a questão do significado e alcance da atuação dos organismos internacionais sobre as condições de vida e saúde das populações dos países pobres, objeto de muitos estudos, mas tão somente apontou-se a sua participação na difusão das metodologias de avaliação. Além disso, é importante salientar que também para os países desenvolvidos, onde se originaram as propostas e práticas de avaliação analisadas nesse texto, não foi desenvolvida ainda uma “avaliação” mais ampla dessas práticas de avaliação desenvolvidas nesses países, e a abordagem desse tema no presente estudo exigiria um estudo específico, dada a sua complexidade. De uma maneira genérica, é possível afirmar que a participação efetiva dos resultados dos processos de avaliação nas grandes decisões políticas para a área da saúde é muito mais restrita do que o anunciado, havendo relativo impacto desses processos em um nível mais específico ou local (Freemantle,²⁴ 1995; Hartz,²⁹ 1998).

No Brasil, não existiram, nem existem, os profissionais e as práticas institucionais tal qual apontadas para os países desenvolvidos e, principalmente, para os Estados Unidos. No entanto, ocorreu também no País grande expansão de atenção à saúde, oferta de novas

tecnologias, desenvolvimento de modelos assistenciais diversos e crescimento da importância política e econômica da atenção à saúde. Condições favoráveis para que propostas de racionalização das decisões e das práticas encontrassem espaço. Contudo, as condições econômicas e sociais, as práticas políticas, a organização profissional, as culturas organizacionais são muito diferentes, o que levou a uma incorporação muito distinta dos diversos tipos de avaliação.

É possível dizer que as avaliações e alternativas metodológicas diversas fazem parte hoje das preocupações de gestores do setor público e privado, mas participam de forma ainda muito incipiente e marginal nos processos de decisão. No setor público, pela atuação da saúde coletiva, há maior interesse pelas avaliações de programas e, até agora, um mais discreto interesse pelas avaliações de qualidade e tecnológicas – a criação de uma agência de vigilância sanitária nos moldes do “Federal Drug Administration” americano poderá alterar esse panorama (Stenzel,⁵⁸ 1996 e os números temáticos de revistas^{4,23,49}). No setor privado, observa-se interesse maior pelas propostas de gestão e garantia da qualidade, além de existirem sistemas de acreditação e monitoramento de hospitais em atividade (Malik,³⁴ 1992; Malik,³⁵ 1996, Novaes & Paganini,⁴⁰ 1994), que poderão ser incorporados pelo setor público hospitalar. O que se constata, no entanto, é um certo recuo na visibilidade dessa questão, sem que tenha ocorrido uma efetiva institucionalização de práticas de avaliação no setor.

REFERÊNCIAS

1. [AHCPR] Agency for Health Care Policy and Research. Translating evidence into practice 1997: conference summary. <http://www.ahcpr.gov/clinic/trip1997/> (Available in 17-11-1998).
2. Aguilar MJ, Ander-Egg E. *Avaliação de serviços e programas sociais*. Petrópolis: Vozes;1994.
3. Altman D. The scandal of poor medical research. *BMJ* 1994;308:283-4.
4. Avaliação tecnológica em saúde. *Divulg Saúde Debate* 1991;(3).
5. Banta HD, Luce BR. *Health care technology and its assessment: an international perspective*. Oxford: Oxford University Press; 1993.
6. Banta HD, Perry S. A History of ISTAHC. A personal perspective on its first 10 years. *Int J Technol Assess Health Care* 1997;13:430-53.
7. Banta HD, Oortwijn WJ, van Beekman WT. *The organization of health care technology assessment in the Netherlands*. The Hague: Rathenau Institute; 1995.
8. Battista RN, Hodge MJ. The development of health care technology assessment: an international perspective. *Int J Technol Assess Health Care* 1995;11:287-300.
9. Bernstein RJ. *Beyond objectivism and relativism: science, hermeneutics and praxis*. Philadelphia: University of Philadelphia Press; 1983.
10. Cabatoff KA. Getting on and off the policy agenda: a dualistic theory of program evaluation utilization. *Can J Program Eval* 1996;11:35-60.
11. Chalmers A. *A fabricação da ciência*. São Paulo: Editora UNESP; 1994.
12. Chen HT. *Theory-driven evaluations*. Newbury Park: Sage Publications; 1990.
13. Contandriopoulos A-P, Champagne F, Denis JF, Pineault R. A avaliação na área da saúde: conceitos e métodos. In: Hartz ZMA organizador. *Avaliação em saúde*. Rio de Janeiro: Editora Fiocruz; 1997. p. 29-48.
14. Denzin NK, Lincoln YS. *Handbook of qualitative research*. Thousand Oaks: Sage Publications; 1994.

15. Donabedian A. Continuity and change in the quest for quality. *Clin Perform Qual Health Care* 1996;1:9-16.
16. Donabedian A. Criteria and standards for quality assessment and monitoring. *Qual Rev Bull* 1986;12:99-108.
17. Donabedian A. The assessment of technology and quality. *Int J Technol Assess Health Care* 1988;4:487-96.
18. Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260:1743-8.
19. Donabedian A. The role of outcomes in quality assessment and assurance. *Qual Rev Bull* 1992;18:356-60.
20. Donabedian A. The seven pillars of quality. *Arch Pathol Lab Med* 1990;114:1115-8.
21. Drummond MF, O'Brien BE, Stoddart G, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 1997.
22. Elixhauser A, Halpern M, Schmier J, Luce BR. Health care cost-benefit and cost-effectiveness analysis: an update. *Med Care* 1998; 36 (suppl 5):MS1-MS9, MS18-147.
23. Epidemiologia e avaliação de serviços de saúde. *Cad Saúde Pública* 1996;12(supl 2).
24. Freemantle N. Dealing with uncertainty: will science solve the problems of resource allocation in the U.K. NHS? *Soc Sci Med* 1995;40:1365-70.
25. Glymour C. *Thinking things through: an introduction to philosophical issues and achievements*. Cambridge: MIT Press; 1992.
26. Goodman C. It's time to rethink health care technology assessment. *Int J Technol Assess Health Care* 1992;8:335-58.
27. Greene JC. Qualitative program evaluation: practice and promise. In: *Handbook of qualitative research*. Denzin NK, Lincoln YS, editors. Thousand Oaks: Sage Publications; 1994. p. 531-44.
28. Guba EG, Lincoln YS. *Fourth generation evaluation*. Newbury Park: Sage Publications; 1989.
29. Hartz ZMA. Institutionalizing the evaluation of health programs and policies in France: cuisine internationale over fast food and sur mesure over ready-made. Rio de Janeiro. *Cad Saúde Pública* 1999;15:229-59.
30. Hartz ZMA, organizador. *Avaliação em saúde: dos modelos conceituais à prática na análise da implantação de programas*. Rio de Janeiro: Editora Fiocruz; 1997.
31. House ER. *Evaluating with validity*. Beverly Hill: Sage Publications; 1980.
32. Ishikawa K. *Controle da qualidade total à maneira japonesa*. Rio de Janeiro: Campus; 1993.
33. Luce BR, Brown R E. The use of technology assessment by hospitals, health maintenance organizations and third party payers in the United States. *Int J Technol Assess Health Care* 1995;11:79-92.
34. Malik AM. Desenvolvimento de recursos humanos, gerência de qualidade e cultura das organizações de saúde. *Rev Adm Empres* 1992;32:32-41.
35. Malik AM. Qualidade em serviços de saúde nos setores público e privado. *Cad Fundap* 1996;19:7-24.
36. Miles MB, Huberman AM. *Qualitative data analysis*. Thousand Oaks: Sage Publications; 1994.
37. Minayo MCS. *O desafio do conhecimento: pesquisa qualitativa em saúde*. São Paulo: Hucitec/Abrasco; 1992.
38. Mohr LB. *Impact analysis for program evaluation*. Thousand Oaks: Sage Publications; 1995.
39. Nogueira RP. *Perspectivas da qualidade em saúde*. Rio de Janeiro: Qualitymark; 1994.
40. Novaes HM, Paganini JM. *Padrões e indicadores de qualidade para hospitais (Brasil)*. Washington (DC): OPAS; 1994.
41. Novaes HMD. *A puericultura em questão* [Dissertação]. São Paulo: Faculdade de Medicina da Universidade de São Paulo; 1979.
42. Novaes HMD. Epidemiologia e avaliação em serviços de atenção médica: novas tendências de pesquisa. *Cad Saúde Pública* 1996;12(Supl 2):7-12.
43. Novaes RL. Sobre a técnica. *Hist Ciênc Saúde* 1996;3:24-49.
44. Panerai RB, Mohr JP. *Health technology assessment methodologies for developing countries*. Washington (DC): PAHO; 1989.
45. Parrochia D. *Sciences exactes et sciences de l'homme: les grandes étapes*. Paris: Ellipses; 1997.
46. Patton MQ. *Utilization-focused evaluation*. Thousand Oaks: Sage Publications; 1997.
47. Perry S, Gardner E, Thamer M. The status of health technology assessment worldwide. *Int J Technol Assess Health Care* 1997;13:81-98.
48. Portocarrero V. *Filosofia, história e sociologia das ciências*. Rio de Janeiro: Editora Fiocruz; 1994.
49. Qualidade em saúde. *Cad Fundap* 1996;(19).
50. Quine WV, Ullian JS. *The web of belief*. New York: Random House; 1970.
51. Rivera FJU. Planejamento estratégico-situacional ou controle de qualidade total em saúde? Um contraponto teórico-metodológico. *Cad Fundap* 1996;19:25-46.
52. Rothman K. *Causal inference*. Chestnut Hill: Epidemiology Resources; 1988.
53. Rossi PH, Freeman HE. *Evaluation: a systematic approach*. Thousand Oaks: Sage Publications; 1993.
54. Schoenbaum SC, editor. *Using clinical practice guidelines to evaluate quality of care*. Washington (DC): US Department of Health and Human Services. Agency for Health Care Policy and Research; 1995. v. 1.

55. Scrivens E. *Accreditation*. Buckingham: Open University Press; 1995.
56. Shortell SM, Richardson WC. *Health program evaluation*. St Louis: Mosby; 1978.
57. Silva LMV, Formigli VLA. Avaliação em saúde: limites e perspectivas. *Cad Saúde Pública* 1994;10:80-91.
58. Stenzel ACB. *A temática da avaliação no campo de saúde coletiva: uma bibliografia comentada* [Dissertação]. Campinas: Universidade Estadual de Campinas; 1996.
59. [UNICEF] United Nations Children's Fund. *An UNICEF guide for monitoring and evaluation*. <http://unicef.org/otcgi/llscgi60.exe>, 1998.
60. Witkowski N. *L'état des sciences*. Paris: La Découverte; 1991.
61. World Health Organization. *Rapid evaluation method: guidelines for maternal and child health, family planning and other health services*. Geneva: WHO; 1993.
62. Yates BT. *Analyzing costs, procedures, processes and outcome in human services*. Thousand Oaks: Sage Publications; 1996.